On the reversibility of the observed process of three-state hidden Markov model

CHEN Yong

School of Mathematics and Computing Science, Hunan University of Science and Technology, Xiangtan, Hunan, 411201, P.R.China. chenyong77@gmail.com

Abstract

For the continuous-time and the discrete-time three-state hidden Markov model, the flux of the likelihood function up to 3-dimension of the observed process is shown explicitly. As an application, the sufficient and necessary condition of the reversibility of the observed process is shown.

keywords: hidden Markov models; likelihood function(joint probability distribution); reversibility.

MSC numbers: 60J27; 60J99; 60K99

1 Introduction

There has been a large amount of literature published on the time reversibility in probability, such as References [1, 2, 3, 4, 5, 6, 7, 8], which are mainly about the Markov processes and the semi-Markov processes (or Markov renewal processes). In the case of a Markov process with finite state (discrete time or continuous time), Kolmogorov's criterion for time reversibility is well-known. In the References [9, 10, 11, 12, 13], they examined time reversibility in the context of a univariate stationary linear time series (Gaussian or non-Gaussian) and of multivariate linear processes.

In Reference[14], for the hidden Markov model, it is shown that the reversibility of the observed process is not equivalent to that of the underlying Markov chain, i.e., if the underlying Markov chain is reversible, then the observed process is reversible too, however, if the Markov chain is irreversible, then the observed process is either reversible or irreversible. In Reference [15], the necessary and sufficient conditions for reversibility of hidden Markov chains on general (countable) spaces are obtained, however, the reversibility therein is concerning the complete process, i.e., the bivariate stochastic process containing both the underlying process and the observed process. That is to say, the above two types of reversibility of the hidden Makov model are different completely.

In the present paper, for continuous-time three-state Markov processes, we calculate the flux of the likelihood function (joint probability distribution) of the observed process. As an application, the sufficient and necessary condition of the reversibility of the three-state hidden Markov model is shown (in the sense of [14], not in the sense of [15]). In fact, besides the reversibility of the underlying Markov process, the reversibility of the observed process is distinguished by whether the state-dependent probability matrix is regular (Definition 3.2, The-

orem 3.4). We illustrate by an example that one cannot detect irreversibility in some cases by comparing directional moments like that used in [14, p104].

We have also investigated the discrete-time three-state hidden Markov model. Since the method is similar to the continuous-time case, we list the conclusions in Appendix (Section 4) and omit most of the proofs. For the discrete-time case, the reversibility is also related to whether zero is an eigenvalue of the 1-step transition probability matrix (Proposition 4.6). Here we see a difference between discrete-time and continuous-time hidden Markov model.

The reversibility of the hidden Markov model may be of interest in some biological studies. An approach to modelling the DNA sequence is to use a hidden Markov model; see, for example, Reference [16, 17]. Since DNA sequences have directions, we should rule out the reversible hidden Markov model.

2 The flux of the likelihood function

Let $\{S_t : t \in \mathbb{R}^+\}$ be the observed process with state space $\mathcal{S} = \{0, 1, 2, \cdots, K-1\}$.

Definition 2.1. The n-dimension likelihood function of $\{S_t : t \in \mathbb{R}^+\}$ is defined as $\Pr(S_{t_1} = s_1, S_{t_2} = s_2, \cdots, S_{t_n} = s_n)$, where $n \in \mathbb{N}$ and $0 \leqslant t_1 \leqslant t_2 \leqslant \cdots \leqslant t_n$. The flux of the likelihood function of $\{S_t : t \in \mathbb{R}^+\}$ is defined as

$$\Pr(S_{t_1} = s_1, S_{t_2} = s_2, \cdots, S_{t_n} = s_n) - \Pr(S_{t_1^-} = s_1, S_{t_2^-} = s_2, \cdots, S_{t_n^-} = s_n), \quad (1)$$

where $t_k^- = t_1 + t_n - t_k$.

Let $\{C_t : t \in \mathbb{R}^+\}$ be an irreducible three-state Markov process with the transition rate matrix Q, and the stationary distribution $\mu = (\mu_1, \mu_2, \mu_3)$, where $\mu_1 + \mu_2 + \mu_3 = 1$, $\mu_i > 0$.

$$\mathsf{Q} = \begin{bmatrix} -a_1 & a_2 & a_3 \\ b_1 & -b_2 & b_3 \\ c_1 & c_2 & -c_3 \end{bmatrix}, \quad (2)$$

where $a_1 = a_2 + a_3$, $b_2 = b_1 + b_3$, $c_3 = c_1 + c_2$, $a_i, b_i, c_i \geqslant 0$, $i = 1, 2, 3$, and $a_1 b_2 c_3$, $b_1 + c_1$, $a_2 + c_2$, $a_3 + b_3 > 0$ (i.e. irreducible). By the stationarity, it is clear that the transition rate flux of $\{C_t : t \in \mathbb{R}^+\}$ is

$$\mu_1 a_2 - \mu_2 b_1 = \mu_2 b_3 - \mu_3 c_2 = \mu_3 c_1 - \mu_1 a_3. \quad (3)$$

Let $\nu = \mu_1 a_2 - \mu_2 b_1$. And the eigen-equation of Q is

$$\lambda(\lambda^2 + \alpha\lambda + \beta) = 0. \quad (4)$$

Denote by $-\lambda_1$, $-\lambda_2$ the nonzero eigenvalues of Q. Let $\triangle = \alpha^2 - 4\beta$.

Similar to the denotation of Reference [18], let S_1^j and T_1^j denote the sequence from 1 to j of the observed states and observation times. The Markov assumption for the hidden process is given by

$$\begin{aligned} &\Pr[C(t_j) \,|\, C(t_1), C(t_2), \ldots, C(t_{j-1}), S_1^{j-1}, T_1^j = t_1^j] \\ &= \Pr[C(t_j) \,|\, C(t_{j-1}), T_{j-1}^j = t_{j-1}^j] \\ &= \mathsf{P}_{c_{j-1}, c_j}(t_j - t_{j-1}), \end{aligned} \quad (5)$$

where the quantity $\mathsf{P}_{c_{j-1},c_j}(t_j - t_{j-1})$ denotes the transition probability of occupying sate c_j at time $T_j = t_j$ given that the process was in state c_{j-1} at t_{j-1}. As indicated by the last equality, the transition probabilities of the process are assumed to be time homogeneous. We also assume that, conditional on the state of the hidden process at time t_j, an observation S_j is independent of all previous observations and the hidden process prior to time t_j:

$$\begin{aligned}&\Pr[S_j \mid C(t_1), C(t_2), \ldots, C(t_j), S_1^{j-1}, T_1^j = t_1^j]\\&= \Pr[S_j \mid C(t_j), T_j = t_j]\\&= \pi(s_j \mid c_j).\end{aligned} \tag{6}$$

Let the 'state-dependent probability' (i.e., emission probability, signal probability) matrix be $\Pi = (\pi(k \mid i))$, $i = 1, 2, 3$; $k = 0, 1, 2, \ldots, K-1$ (i.e., a $3 \times K$ matrix). Note that the rows of Π must sum to 1 [1]. Let φ_k be the k-column of Π and $\wedge_k = \mathrm{diag}\,\{\varphi_k\}$, $k = 0, 1, 2, \ldots, K-1$.

Proposition 2.2. *The flux of the 2-dimension likelihood function is when* $t > 0$*,*

$$\Pr\{S_0 = i, S_t = j\} - \Pr\{S_0 = j, S_t = i\} = \frac{\nu A}{\lambda_1 - \lambda_2}[e^{-\lambda_2 t} - e^{-\lambda_1 t}],$$

where ν *is the transition rate flux,* $A = (y_2 - x_2)(x_1 - z_1) - (x_2 - z_2)(y_1 - x_1)$*,* $(x_1, y_1, z_1)' = \varphi_i$*, and* $(x_2, y_2, z_2)' = \varphi_j$*.*

Corollary 2.3. *If the rank of the state-dependent probability is* 1 *or* 2*, then* $\Pr\{S_0 = i, S_t = j\} - \Pr\{S_0 = j, S_t = i\} = 0$.

Theorem 2.4. *The flux of the following 3-dimension likelihood function is when* r, $t > 0$*,*

$$\begin{aligned}&\Pr\{S_0 = S_r = S_{r+t} = i\} - \Pr\{S_0 = S_t = S_{t+r} = i\}\\&= \tfrac{\nu D}{\lambda_1 - \lambda_2}(e^{-\lambda_2 r - \lambda_1 t} - e^{-\lambda_2 t - \lambda_1 r}),\end{aligned}$$

where ν *is the transition rate flux,* $D = (x - y)(y - z)(z - x)$*,* $(x, y, z)' = \varphi_i$*.*

Proofs of Proposition 2.2, Corollary 2.3, Theorem 2.4 are presented in Subsection 2.1.

2.1 Proofs

Let $\mathsf{U} = \mathrm{diag}\,\{\mu_1, \mu_2, \mu_3\}$. Then

$$\mathsf{UQ} - \mathsf{Q'U} = \nu \begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix} \tag{7}$$

Let $\vec{e} = (1, 1, 1)'$ and the matrix $\mathsf{L} = \vec{e}\mu$.

Lemma 2.5. *If* $\triangle \neq 0$*, then for* $t \in \mathbb{R}^+$*, the* t*-step transition probability matrix is*

$$\mathsf{P}(t) = g_t\mathsf{L} + d_t\mathsf{Q} + f_t\mathsf{I}, \tag{8}$$

[1] The state-dependent probability here is the transpose matrix of that in Reference [14].

where

$$\begin{aligned} d_t &= \tfrac{e^{-\lambda_2 t} - e^{-\lambda_1 t}}{\lambda_1 - \lambda_2}, \\ f_t &= \tfrac{\lambda_1 e^{-\lambda_2 t} - \lambda_2 e^{-\lambda_1 t}}{\lambda_1 - \lambda_2}, \\ g_t &= 1 - f_t. \end{aligned} \tag{9}$$

It is Proposition 4.3 of Reference [19]. The reader can also refer to Theorem 14.9 of Reference [20]. To write it in terms of function with matrix coefficients is the key to the results in the present paper.

Remark 1. Fix the value of λ_1, and let $\lambda_2 \to \lambda_1$, then one has

$$\begin{aligned} d_t &= t\, e^{-\lambda_1 t}, \\ f_t &= (1 + \lambda_1 t) e^{-\lambda_1 t}, \\ g_t &= 1 - f_t. \end{aligned} \tag{10}$$

It is exactly the t-step transition probability matrix in the case $\triangle = 0$, please refer to [19, 20] and the references therein. That is to say, Eq.(10) is the same as Eq.(9) in the sense of limit. We therefore do not distinguish whether $\triangle = 0$ or not for all the subsequent formulas.

Proof of Proposition 2.2. Since $\mathsf{L} = \vec{e}\mu$, we have

$$\mu \wedge_i \mathsf{L} \wedge_j \vec{e} = (\mu \wedge_i \vec{e})(\mu \wedge_j \vec{e}) = (\mu \wedge_j \vec{e})(\mu \wedge_i \vec{e}) = \mu \wedge_j \mathsf{L} \wedge_i \vec{e}.$$

Note that $\wedge_i \wedge_j = \wedge_j \wedge_i$. By Eq.(2.27) of Reference [14] and Lemma 2.5, we have

$$\begin{aligned} & \Pr\{S_0 = i, S_t = j\} - \Pr\{S_0 = j, S_t = i\} \\ = \;& \mu \wedge_i \mathsf{P}(t) \wedge_j \vec{e} - \mu \wedge_j \mathsf{P}(t) \wedge_i \vec{e} \\ = \;& \mu \wedge_i (g_t \mathsf{L} + d_t \mathsf{Q} + f_t \mathsf{I}) \wedge_j \vec{e} - \mu \wedge_j (g_t \mathsf{L} + d_t \mathsf{Q} + f_t \mathsf{I}) \wedge_i \vec{e} \\ = \;& g_t(\mu \wedge_i \mathsf{L} \wedge_j \vec{e} - \mu \wedge_j \mathsf{L} \wedge_i \vec{e}) + d_t(\mu \wedge_i \mathsf{Q} \wedge_j \vec{e} - \mu \wedge_j \mathsf{Q} \wedge_i \vec{e}) + f_t(\mu \wedge_i \wedge_j \vec{e} - \mu \wedge_j \wedge_i \vec{e}) \\ = \;& d_t(\mu \wedge_i \mathsf{Q} \wedge_j \vec{e} - \mu \wedge_j \mathsf{Q} \wedge_i \vec{e}). \end{aligned}$$

Since $\mu \wedge_i \mathsf{Q} \wedge_j \vec{e} = \varphi_i' \mathsf{U} \mathsf{Q} \varphi_j$, we have

$$\mu \wedge_j \mathsf{Q} \wedge_i \vec{e} = \varphi_j' \mathsf{U} \mathsf{Q} \varphi_i = (\varphi_j' \mathsf{U} \mathsf{Q} \varphi_i)' = \varphi_i' \mathsf{Q}' \mathsf{U} \varphi_j.$$

By Eq.(7), we have

$$\mu \wedge_i \mathsf{Q} \wedge_j \vec{e} - \mu \wedge_j \mathsf{Q} \wedge_i \vec{e} = \varphi_i' (\mathsf{U}\mathsf{Q} - \mathsf{Q}'\mathsf{U}) \varphi_j = \nu A. \tag{11}$$

This ends the proof. □

Proof of Corollary 2.3. Note that $A = \det\{\mathsf{H}\}$, where "det" is the determinant function, and $\mathsf{H} = [\vec{e}, \varphi_i, \varphi_j]$ is a 3×3 matrix. If the rank of the state-dependent probability is 1, then φ_i, φ_j are linear dependent and we obtain that $A = 0$. If the rank of the state-dependent probability is 2, and if φ_i, φ_j are linear independent, then they are one base of $\{\varphi_k, k = 0, 1, 2, \ldots, K-1\}$. Note that $\vec{e} = \sum_{k=0}^{K-1} \varphi_i$. Then φ_i, φ_j, $\vec{e}$ are linear dependent. Thus $A = 0$. This ends the proof. □

Proof of Theorem 2.4. Since $\mathsf{L} = \vec{e}\mu$, we have

$$\mu \wedge_i \mathsf{L} \wedge_i \mathsf{Q} \wedge_i \vec{e} = (\mu \wedge_i \vec{e})(\mu \wedge_i \mathsf{Q} \wedge_i \vec{e}) = (\mu \wedge_i \mathsf{Q} \wedge_i \vec{e})(\mu \wedge_i \vec{e}) = \mu \wedge_i \mathsf{Q} \wedge_i \mathsf{L} \wedge_i \vec{e}.$$

Similar to Proposition 2.2, we obtain

$$
\begin{aligned}
& \Pr\{S_0 = S_r = S_{r+t} = i\} - \Pr\{S_0 = S_t = S_{t+r} = i\} \\
= \;& \mu\wedge_i \mathsf{P}(r)\wedge_i \mathsf{P}(t)\wedge_i \vec{e} - \mu\wedge_i \mathsf{P}(t)\wedge_i \mathsf{P}(r)\wedge_i \vec{e} \\
= \;& \mu\wedge_i (g_r\mathsf{L} + d_r\mathsf{Q} + f_r\mathsf{I})\wedge_i (g_t\mathsf{L} + d_t\mathsf{Q} + f_t\mathsf{I})\wedge_i \vec{e} \\
- \;& \mu\wedge_i (g_t\mathsf{L} + d_t\mathsf{Q} + f_t\mathsf{I})\wedge_i (g_r\mathsf{L} + d_r\mathsf{Q} + f_r\mathsf{I})\wedge_i \vec{e} \\
= \;& (g_r d_t - g_t d_r)(\mu\wedge_i \mathsf{L}\wedge_i \mathsf{Q}\wedge_i \vec{e} - \mu\wedge_i \mathsf{Q}\wedge_i \mathsf{L}\wedge_i \vec{e}) + (g_r f_t - g_t f_r)(\mu\wedge_i \mathsf{L}\wedge_i^2 \vec{e} - \mu\wedge_i^2 \mathsf{L}\wedge_i \vec{e}) \\
+ \;& (d_r f_t - d_t f_r)(\mu\wedge_i \mathsf{Q}\wedge_i^2 \vec{e} - \mu\wedge_i^2 \mathsf{Q}\wedge_i \vec{e}) \\
= \;& (d_r f_t - d_t f_r)(\mu\wedge_i \mathsf{Q}\wedge_i^2 \vec{e} - \mu\wedge_i^2 \mathsf{Q}\wedge_i \vec{e}).
\end{aligned}
$$

It follows from Lemma 2.5 that

$$d_r f_t - d_t f_r = \frac{1}{\lambda_1 - \lambda_2}[e^{-\lambda_2 r - \lambda_1 t} - e^{-\lambda_2 t - \lambda_1 r}].$$

Let $\psi = (x^2, y^2, z^2)'$. Similar to Eq.(11), we have

$$
\begin{aligned}
& \mu\wedge_i \mathsf{Q}\wedge_i^2 \vec{e} - \mu\wedge_i^2 \mathsf{Q}\wedge_i \vec{e} \\
& = \varphi_i'(\mathsf{UQ} - \mathsf{Q'U})\psi \\
& = \nu[x(y^2 - z^2) + y(z^2 - x^2) + z(x^2 - y^2)] \\
& = \nu D.
\end{aligned} \tag{12}
$$

This ends the proof. □

3 The reversibility of the observed process

The observed process is the same as in Section 2.

Definition 3.1. The observed process is said to be reversible if its finite-dimensional distributions are invariant under reversal of time, i.e., the flux of the likelihood function vanishes,

$$\Pr(S_{t_1} = s_1,\, S_{t_2} = s_2,\, \cdots,\, S_{t_n} = s_n) - \Pr(S_{t_1^-} = s_1,\, S_{t_2^-} = s_2,\, \cdots,\, S_{t_n^-} = s_n) = 0,$$

where $t_k^- = t_1 + t_n - t_k$, for all positive integers n and all $0 \leqslant t_1 \leqslant t_2 \leqslant \cdots \leqslant t_n$.

Definition 3.2. Two rows of a matrix are said to be equal if they are two equal vectors. If any two rows of the state-dependent probability matrix Π are not equal, we say that Π is regular. Otherwise, we say that Π is singular, i.e., there are at least two undistinguishable states among the three hidden states by means of observation.

Theorem 3.3. *If the underlying Markov process is reversible, then the observed process is reversible too.*

Although the proof in Reference [14, P102-103] is about the discrete-time hidden Markov model, it is still valid for the continuous-time one and is ignored here.

Theorem 3.4. *The observed process of the continuous-time three-state hidden Makov model is irreversible, if and only if the underlying Markov process is irreversible and the state-dependent probability matrix is regular.*

Proof of Theorem 3.4 is presented in Section 3.2.

If Π is regular, the rank of Π is 3 or 2. If Π is singular, the rank of Π is 2 or 1. That is to say, the rank of Π is involved in the reversibility of the observed processes.

By Reference [21], the hard limiting (or clipping) transformation is very useful from a practical viewpoint, and the rhythm inherited in the binary series carries a great deal of information about the original series. If we maintain the regularity condition of Π when clipping the observed process of the hidden Markov model, it preserves the time-reversibility property by the last theorem (e.g., Example 2).

The time reversibility of high-order hidden Markov models (e.g., more than four-state) is difficult to be solved completely. We can only find some simple sufficient conditions of the irreversibility, for example, the underlying Markov process is irreversible and the rank of the state-dependent probability matrix is equal to the number of states of the underlying Markov process (similar to Proposition 3.5).

Remark 2. Let the complete process be $\{X_t = (C_t,\, S_t),\, t \geqslant 0\}$. Clearly, it is still be a finite-state Markov process, please refer to [22]. Similar to Theorem 2.1 of Reference [15], by Kolmogorov's criterion, we can show that the complete process is reversible if and only if the underlying process is reversible. That is to say, there are two different types of reversibility of the hidden Makov model.

3.1 Applications

Example 1. The deterministic function of a Markov process is a special case of hidden Markov model. Let $f = (f_1,\, f_2,\, f_3)'$ be a function defined on the state space. If $f = (1, 1, 0)'$ or $f = (1, 0, 0)'$ like that used in Reference [23], then the state-dependent probability matrices are respectively

$$\Pi_1 = \begin{bmatrix} 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad \Pi_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 0 \end{bmatrix}.$$

Since the state-dependent probability matrices are singular, the observed processes is reversible by Theorem 3.4.

Example 2. Suppose $\{C_t,\, t \geqslant 0\}$ be the irreversible Markov process with transition rate matrix

$$Q = \begin{bmatrix} -2/3 & 1/3 & 1/3 \\ 2/3 & -1 & 1/3 \\ 1/2 & 1/2 & -1 \end{bmatrix}. \tag{13}$$

Let $\{S_t\}^2$, $\{\xi_t\}$, $\{\eta_t\}$ be three observed processes with state-dependent probability matrices respectively

$$\Pi_1 = \begin{bmatrix} 1 & 0 & 0 \\ 1/4 & 1/2 & 1/4 \\ 0 & 0 & 1 \end{bmatrix}, \quad \Pi_2 = \begin{bmatrix} 1 & 0 & 0 \\ 1/4 & 1/2 & 1/4 \\ 1/2 & 1/3 & 1/6 \end{bmatrix}, \quad \Pi_3 = \begin{bmatrix} 1 & 0 \\ 1/4 & 3/4 \\ 0 & 1 \end{bmatrix}.$$

Since all the state-dependent probability matrices are regular, the observed processes are irreversible by Theorem 3.4. $\{\eta_t\}$ is clipped from $\{S_t\}$ and preserves

[2] $\{S_t\}$ comes from the example in Reference [14, p105].

irreversible. Since both the rank of Π_2 and Π_3 are 2, by Corollary 2.3, we have that

$$\begin{aligned}\Pr\{\xi_t = i, \xi_{t+r} = j\} &= \Pr\{\xi_t = j, \xi_{t+r} = i\},\\ \Pr\{\eta_t = i, \eta_{t+r} = j\} &= \Pr\{\eta_t = j, \eta_{t+r} = i\}.\end{aligned}$$

In the case, one cannot detect irreversibility by comparing directional moments $\mathbb{E}(\xi_t\xi_{t+r}^n)$ and $\mathbb{E}(\xi_t^n\xi_{t+r})$ with $n \in \mathbb{N}$ like that used in Reference [14, p104].

3.2 Proof

Let $\vec{e}_1 = (1,0,0)', \vec{e}_2 = (0,1,0)'$, $\vec{e}_3 = (0,0,1)'$, $\mathsf{E}_i = \operatorname{diag}\{\vec{e}_i\}$, $\wedge_k = \operatorname{diag}\{\varphi_k\}$.

Proposition 3.5. *If the underlying Markov process is irreversible and the rank of the state-dependent probability is 3, then the observed process is irreversible.*

Proof. Since the rank of the state-dependent probability is 3, we can choose a base of $\mathbb{R}^3$, without loss generality, to be $\varphi_0, \varphi_1, \varphi_2$. Then $\vec{e}_1 = \sum_{i=0}^2 x_i\varphi_i$, $\vec{e}_2 = \sum_{j=0}^2 y_j\varphi_j$, and $\mathsf{E}_1 = \sum_{i=0}^2 \wedge_i x_i$, $\mathsf{E}_2 = \sum_{j=0}^2 \wedge_j y_j$.

Suppose on the contrary that the observed process is reversible. When $t > 0$,

$$\begin{aligned}0 &= \Pr\{S_0 = i, S_t = j\} - \Pr\{S_0 = j, S_t = i\}\\ &= \mu\wedge_i\mathsf{P}(t)\wedge_j\vec{e} - \mu\wedge_j\mathsf{P}(t)\wedge_i\vec{e}, \quad \text{where} \quad i,j = 0,1,2.\end{aligned}$$

Thus

$$\begin{aligned}\mu_1\mathsf{P}_{12}(t) - \mu_2\mathsf{P}_{21}(t) &= \mu\mathsf{E}_1\mathsf{P}(t)\mathsf{E}_2\vec{e} - \mu\mathsf{E}_2\mathsf{P}(t)\mathsf{E}_1\vec{e}\\ &= \sum_{i,j=0}^{2} x_iy_j\mu\wedge_i\mathsf{P}(t)\wedge_j\vec{e} - \sum_{i,j=0}^{2} y_jx_i\mu\wedge_j\mathsf{P}(t)\wedge_i\vec{e}\\ &= \sum_{i,j=0}^{2} x_iy_j[\mu\wedge_i\mathsf{P}\wedge_j\vec{e} - \mu\wedge_j\mathsf{P}\wedge_i\vec{e}]\\ &= 0\end{aligned}$$

Note that

$$\lim_{t\to 0+}\frac{\mathsf{P}(t) - \mathsf{I}}{t} = \mathsf{Q}.$$

Then $\nu = \mu_1 a_2 - \mu_2 b_1 = 0$. However, the underlying Markov process is irreversible, i.e., $\nu \neq 0$, a contradiction. □

Lemma 3.6. *If the rank of* Π *is* 2*, then all its columns* $\{\varphi_i, i = 0,1,2,\ldots,K-1\}$ *are the linear combination of its certain column and* $\vec{e} = (1,1,1)'$.

Proof. If the rank is 2, then $K \geqslant 2$. Without loss generality, let $\{\varphi_1,\, \varphi_2\}$ be one base of $\{\varphi_i, i = 0,1,2,\ldots,K-1\}$ with $\varphi_1 \neq c\,\vec{e}$, where $c \in \mathbb{R}$. Note that $\vec{e} = \sum_{i=0}^{K-1}\varphi_i = x\varphi_1 + y\varphi_2$, where $y \neq 0$, $x \in \mathbb{R}$, thus $\varphi_2 = (\vec{e} - x\varphi_1)/y$. Since $\{\varphi_1,\, \varphi_2\}$ is one base, all $\{\varphi_i, i = 0,1,2,\ldots,K-1\}$ are the linear combination of $\{\varphi_1,\, \vec{e}\}$. □

Without loss generality, if the rank of Π is 2, let $\{\varphi_i, i = 0,1,2,\ldots,K-1\}$ be the linear combination of φ_1 and $\vec{e}$. Let $\wedge = \operatorname{diag}\{\varphi_1\}$.

Proposition 3.7. *Suppose that the underlying Markov process is irreversible. If the rank of* Π *is* 2*, and if any two rows of* Π *are not equal, then the observed process is irreversible.*

Proof. We claim that $\varphi_1 = (x, y, z)'$ with $x \neq y \neq z$. Without loss of generality, suppose on the contrary that $\varphi_1 = (x, x, z)'$, then the first two rows of Π are equal by Lemma 3.6, a contradiction. Thus

$$D = (x - y)(y - z)(z - x) \neq 0.$$

Since the underlying Markov process is irreversible, the transition rate flux $\nu \neq 0$. By Theorem 2.4, when $r, t > 0$ and $r \neq t$,

$$\Pr\{S_0 = S_r = S_{r+t} = 1\} - \Pr\{S_0 = S_t = S_{t+r} = 1\} \neq 0,$$

i.e., the observed process is irreversible. $\square$

Proposition 3.8. *If there are two equal rows of* Π*, then the observed process is time reversible.*

Proof. If the rank of Π is 1, i.e., all the three rows of Π are equal, the observed process is in fact identical independent distribution series.

If the rank of Π is 2, without loss of generality, suppose the first and the second row of Π are equal. Then $\varphi_{s_k} = x_k\vec{e} + y_k\vec{e}_3$ and $\wedge_{s_k} = x_k\mathsf{I} + y_k\mathsf{E}_3$. The flux of the likelihood function is

$$\begin{aligned}
& \Pr(S_{t_1} = s_1,\, S_{t_1+t_2} = s_2,\, \cdots,\, S_{t_1+t_2+\cdots+t_r} = s_r) \\
- & \Pr(S_{t_1} = s_r,\, S_{t_1+t_r} = s_{r-1}, \cdots,\, S_{t_1+t_r+\cdots+t_2} = s_1) \\
= & \mu \wedge_{s_1} \mathsf{P}(t_2) \wedge_{s_2} \mathsf{P}(t_3) \cdots \mathsf{P}(t_r) \wedge_{s_r} \vec{e} - \mu \wedge_{s_r} \mathsf{P}(t_r) \wedge_{s_{r-1}} \mathsf{P}(t_{r-1}) \cdots \mathsf{P}(t_2) \wedge_{s_1} \vec{e} \\
= & \mu[x_1\mathsf{I} + y_1\mathsf{E}_3]\mathsf{P}(t_2)[x_2\mathsf{I} + y_2\mathsf{E}_3]\mathsf{P}(t_3) \cdots \mathsf{P}(t_r)[x_r\mathsf{I} + y_r\mathsf{E}_3]\vec{e} \\
- & \mu[x_r\mathsf{I} + y_r\mathsf{E}_3]\mathsf{P}(t_r)[x_{r-1}\mathsf{I} + y_{r-1}\mathsf{E}_3]\mathsf{P}(t_{r-1}) \cdots \mathsf{P}(t_2)[x_1\mathsf{I} + y_1\mathsf{E}_3]\vec{e}.
\end{aligned}$$

Expand the expression, and delete the terms which do not contain E_3. Note that $\mathsf{E}_3^l = \mathsf{E}_3$ for $l \in \mathbb{N}$(i.e., E_3 is projective matrix), $\mathsf{P}(t)\vec{e} = \vec{e}$, $\mu\mathsf{P}(t) = \mu$, and $\mathsf{P}(t)\mathsf{P}(r) = \mathsf{P}(t + r)$ for $t,\, r \in \mathbb{R}^+$. All other terms pairwise satisfy that

$$\begin{aligned}
& \mu \cdots \mathsf{P}(t_i)\mathsf{I} \cdots \mathsf{P}(t_j)\mathsf{E}_3 \cdots \vec{e} - \mu \cdots \mathsf{E}_3\mathsf{P}(t_j) \cdots \mathsf{I}\mathsf{P}(t_i) \cdots \vec{e} \\
= & \mu(\mathsf{E}_3\mathsf{P}(r)\mathsf{E}_3) \cdots (\mathsf{E}_3\mathsf{P}(t)\mathsf{E}_3)\vec{e} - \mu(\mathsf{E}_3\mathsf{P}(t)\mathsf{E}_3) \cdots (\mathsf{E}_3\mathsf{P}(r)\mathsf{E}_3)\vec{e} \\
= & \mu_3\mathsf{P}_{33}(r) \cdots \mathsf{P}_{33}(t) - \mu_3\mathsf{P}_{33}(t) \cdots \mathsf{P}_{33}(r) \\
= & 0.
\end{aligned}$$

This ends the proof. $\square$

Proposition 3.9. *Suppose that the underlying Markov process is irreversible.*

1) *If the state-dependent probability matrix* Π *is singular, then the observed process is reversible.*

2) *If the state-dependent probability matrix* Π *is regular, then the observed process is irreversible.*

Proof. The first case is Proposition 3.8. If Π is regular, then the rank of Π is 3 or 2. Thus the second case is Proposition 3.7 and Proposition 3.5. $\square$

Proof of Theorem 3.4. It can be shown directly by Theorem 3.3 and Proposition 3.9. $\square$

4 Appendix: the discrete-time case

Let $\{S_t : t \in \mathbb{Z}^+\}$ be the observed process with state space $\mathcal{S} = \{0, 1, 2, \cdots, K-1\}$. Let $\{C_t : t \in \mathbb{Z}^+\}$ be an irreducible three-state Markov chain with the 1-step transition probability matrix P and the stationary distribution $\mu = (\mu_1, \mu_2, \mu_3)$, where $\mu_1 + \mu_2 + \mu_3 = 1$, $\mu_i > 0$.

$$\mathsf{P} = \begin{bmatrix} 1-a_2-a_3 & a_2 & a_3 \\ b_1 & 1-b_1-b_3 & b_3 \\ c_1 & c_2 & 1-c_1-c_2 \end{bmatrix} \tag{14}$$

By the stationarity, it is clear that the probability flux is

$$\mu_1 a_2 - \mu_2 b_1 = \mu_2 b_3 - \mu_3 c_2 = \mu_3 c_1 - \mu_1 a_3. \tag{15}$$

Let $\nu = \mu_1 a_2 - \mu_2 b_1$. Let $\mathsf{Q} = \mathsf{P} - \mathsf{I}$, where I is the unit matrix. Denote by $-\lambda_1$, $-\lambda_2$ the nonzero eigenvalues of Q.

Lemma 4.1. *If $\triangle \neq 0$, then for $n \in \mathbb{N}$, the n-step transition probability matrix is*

$$\mathsf{P}^n = g_n \mathsf{L} + d_n \mathsf{Q} + f_n \mathsf{I}, \tag{16}$$

where

$$\begin{aligned} d_n &= \tfrac{(1-\lambda_2)^n - (1-\lambda_1)^n}{\lambda_1 - \lambda_2}, \\ f_n &= \tfrac{\lambda_1(1-\lambda_2)^n - \lambda_2(1-\lambda_1)^n}{\lambda_1 - \lambda_2}, \\ g_n &= 1 - f_n. \end{aligned} \tag{17}$$

Proposition 4.2. *The flux of the 2-dimension likelihood function is when $n \in \mathbb{N}$,*

$$\Pr\{S_0 = i, S_n = j\} - \Pr\{S_0 = j, S_n = i\} = \frac{\nu A}{\lambda_1 - \lambda_2}[(1-\lambda_2)^n - (1-\lambda_1)^n],$$

where ν is the probability flux, $A = (y_2 - x_2)(x_1 - z_1) - (x_2 - z_2)(y_1 - x_1)$, $(x_1, y_1, z_1)' = \varphi_i$, and $(x_2, y_2, z_2)' = \varphi_j$.

Theorem 4.3. *The flux of the following 3-dimension likelihood function is when $n, m \in \mathbb{N}$,*

$$\begin{aligned} &\Pr\{S_0 = S_n = S_{n+m} = i\} - \Pr\{S_0 = S_m = S_{m+n} = i\} \\ &= \tfrac{\nu D}{\lambda_1 - \lambda_2}[(1-\lambda_2)^n (1-\lambda_1)^m - (1-\lambda_1)^n (1-\lambda_2)^m], \end{aligned}$$

where ν is the probability flux, $D = (x-y)(y-z)(z-x)$, $(x, y, z)' = \varphi_i$.

Theorem 4.4. *The three-state hidden Makov models is irreversible, if and only if the underlying Markov chain is irreversible, the state-dependent probability matrix is regular, and one of the following conditions is satisfied:*

1) *the rank of Π is 3,*

2) *the rank of Π is 2, and 0 is not the eigenvalue of P.*

Proofs of Lemma 4.1, Proposition 4.2 and Theorem 4.3 are omitted. Proof of Theorem 4.4 is presented in Subsection 4.1.

4.1 Proofs

Lemma 4.5. *If the rank of* Π *is* 2*, the observed process is reversible if and only if for all* $n, m, \ldots, k \in \mathbb{N}$,

$$\mu\wedge\mathsf{P}^n\wedge\mathsf{P}^m\wedge\cdots\wedge\mathsf{P}^k\wedge\vec{e}-\mu\wedge\mathsf{P}^k\wedge\cdots\wedge\mathsf{P}^m\wedge\mathsf{P}^n\wedge\vec{e}=0, \tag{18}$$

where $\wedge = \mathrm{diag}\,\{\varphi_1\}$.

Proof. The necessity is trivial. We need to prove the sufficiency only. It follows that $\varphi_{s_k} = x_k\vec{e}+y_k\varphi_1$ from Lemma 3.6. Then $\wedge_{s_k} = x_k\mathsf{I}+y_k\wedge$. The flux of the likelihood function is

$$\begin{aligned}
& \Pr(S_1 = s_1,\, S_2 = s_2,\, \cdots,\, S_l = s_l) - \Pr(S_l = s_1,\, S_{l-1} = s_2,\, \cdots,\, S_1 = s_l) \\
= \;& \mu \wedge_{s_1} \mathsf{P} \wedge_{s_2} \mathsf{P}\cdots \wedge_{s_l} \vec{e} - \mu \wedge_{s_l} \mathsf{P} \wedge_{s_{l-1}} \mathsf{P}\cdots \wedge_{s_1} \vec{e} \\
= \;& \mu[x_1\mathsf{I}+y_1\wedge]\mathsf{P}[x_2\mathsf{I}+y_2\wedge]\mathsf{P}\cdots[x_l\mathsf{I}+y_l\wedge]\vec{e} \\
- \;& \mu[x_l\mathsf{I}+y_l\wedge]\mathsf{P}[x_{l-1}\mathsf{I}+y_{l-1}\wedge]\mathsf{P}\cdots[x_1\mathsf{I}+y_1\wedge]\vec{e} \\
= \;& \sum_{\{i_1,i_2,\cdots,i_s\}} x_{i_1}\cdots x_{i_s}y_{j_1}\cdots y_{j_k}\big[\mu\cdots\mathsf{P}\mathsf{I}\cdots\mathsf{P}\wedge\cdots\vec{e}-\mu\cdots\wedge\mathsf{P}\cdots\mathsf{I}\mathsf{P}\cdots\vec{e}\big].
\end{aligned}$$

where $\{i_1, i_2, \cdots, i_s\} \in \{1, 2, \cdots, l\}$ and $\{j_1, j_2, \cdots, j_k\} = \{1, 2, \cdots, l\}\backslash\{i_1, \cdots, i_s\}$. Delete the term which does not contain $\wedge$. Note that $\mu\mathsf{P} = \mu$ and $\mathsf{P}\vec{e} = \vec{e}$.

$$\begin{aligned}
& \mu\cdots\mathsf{P}\mathsf{I}\cdots\mathsf{P}\wedge\cdots\vec{e}-\mu\cdots\wedge\mathsf{P}\cdots\mathsf{I}\mathsf{P}\cdots\vec{e} \\
= \;& \mu\wedge\mathsf{P}^n\wedge\mathsf{P}^m\wedge\cdots\wedge\mathsf{P}^k\wedge\vec{e}-\mu\wedge\mathsf{P}^k\wedge\cdots\wedge\mathsf{P}^m\wedge\mathsf{P}^n\wedge\vec{e} \\
= \;& 0.
\end{aligned}$$

This ends the proof. □

Remark 3. Eq.(18) is equivalent to for all positive integers r and all $0 \leqslant t_1 \leqslant t_2 \leqslant \cdots \leqslant t_r$,

$$\Pr(S_{t_1} = S_{t_2} = \cdots = S_{t_r} = 1) = \Pr(S_{t_1^-} = S_{t_2^-} = \cdots = S_{t_r^-} = 1), \tag{19}$$

where $t_l^- = t_1 + t_r - t_l$.

Proposition 4.6. *If the rank of the state-dependent probability is* 2*, and* 0 *is the eigenvalue of the 1-step transition probability, then the observed process is reversible.*

Proof. Without loss generality, let $1 - \lambda_2 = 0$. By Lemma 2.5, for all $n \in \mathbb{N}$,

$$\mathsf{P}^n = d_n\mathsf{P} + g_n\mathsf{L},$$

where $g_n = 1 - d_n$, $d_n = (1 - \lambda_1)^{n-1}$.

$$\begin{aligned}
& \mu\wedge\mathsf{P}^n\wedge\mathsf{P}^m\wedge\cdots\wedge\mathsf{P}^k\wedge\vec{e}-\mu\wedge\mathsf{P}^k\wedge\cdots\wedge\mathsf{P}^m\wedge\mathsf{P}^n\wedge\vec{e} \\
= \;& \mu\wedge[d_n\mathsf{P}+g_n\mathsf{L}]\wedge[d_m\mathsf{P}+g_m\mathsf{L}]\wedge\cdots\wedge[d_k\mathsf{P}+g_k\mathsf{L}]\wedge\vec{e} \\
- \;& \mu\wedge[d_k\mathsf{P}+g_k\mathsf{L}]\wedge\cdots\wedge[d_m\mathsf{P}+g_m\mathsf{L}]\wedge[d_n\mathsf{P}+g_n\mathsf{L}]\wedge\vec{e} \\
= \;& \sum_{\{i_1,i_2,\cdots,i_s\}} d_{i_1}\cdots d_{i_s}g_{j_1}\cdots g_{j_k}\big[\mu\wedge\cdots\mathsf{P}\wedge\cdots\mathsf{L}\wedge\cdots\wedge\vec{e}-\mu\wedge\cdots\wedge\mathsf{L}\cdots\wedge\mathsf{P}\cdots\wedge\vec{e}\big].
\end{aligned}$$

Delete the term which does not contain L. Note that $\mathsf{L} = \vec{e}\mu$.

$$\begin{aligned} & \mu\wedge\cdots\mathsf{P}\wedge\cdots\mathsf{L}\wedge\cdots\wedge\vec{e} \\ = \; & (\mu\wedge\mathsf{P}\cdots\mathsf{P}\wedge\vec{e})(\mu\wedge\mathsf{P}\cdots\mathsf{P}\wedge\vec{e})\cdots(\mu\wedge\mathsf{P}\cdots\mathsf{P}\wedge\vec{e}) \\ = \; & \mu\wedge\cdots\wedge\mathsf{L}\cdots\wedge\mathsf{P}\cdots\wedge\vec{e}. \end{aligned}$$

This ends the proof by Lemma 4.5. □

Proposition 4.7. *Suppose that the underlying Markov chain is irreversible. Then we have*

1) *if there are two equal rows of* Π, *then the observed process is reversible;*

2) *if the rank of* Π *is* 2,
a) and if 0 *is the eigenvalue of* P, *then the observed process is reversible;*
b) Π *is regular, and if* 0 *is not the eigenvalue of* P, *then the observed process is irreversible;*

3) *if the rank of* Π *is* 3, *then the observed process is irreversible.*

Proof. The first case is similar to Proposition 3.8. The second case is Proposition 4.6 and similar to Proposition 3.7. The third case is similar to Proposition 3.5. □

Proof of Theorem 4.4. It can be shown directly by Theorem 3.3 and Proposition 4.7. □

Acknowledgements

This work is supported by Hunan Provincial Natural Science Foundation of China (No 10JJ6014).

References

[1] Kolmogorov A N: Zur Theorie der Markoffschen Ketten, Math. Ann., 112 (1936), 155-160.

[2] Hostinsky B, Potocek J: Chaines de Markoff inverses, Bull. Intern. Acad. Technique Sci. 36 (1935), 64-67.

[3] Dobrushin R L, Sukhov Y M, Fritz J: A.N. Kolmogorov–the founder of the theory of reversible Markov processes, Uspekhi Mat. Nauk 43:6 (1988), 167-188. Russian Math. Surveys 43:6 , 157-182 (1988).

[4] Qian H, Qian M, Tang X: Thermodynamics of the general diffusion process: time-reversibility and entropy production, J. Stat. Phys. 107, 1129-1141, (2002).

[5] Jiang D-Q, Qian M, Qian M-P: Mathematical Theory of Nonequilibrium Steady States, Berlin: Springer, 2004

[6] Serfozo R F: Reversible Markov processes on general spaces and spatial migration processes, Adv. in Appl. Probab. Volume 37, Number 3 (2005), 801-818.

[7] Ge H, Jiang D-Q, Qian M: Reversibility and entropy production of inhomogeneous Markov chains, J. Appl. Probab. Volume 43, Number 4, 1028-1043, (2006).

[8] Wang H, Qian H: On detailed balance and reversibility of semi-Markov processes and single-molecule enzyme kinetics J. Math. Phys. 48, 013303 (2007).

[9] Weiss G: Time reversibility of linear stochastic processes. J. Appl. Probability, 12, 831-836 (1975).

[10] Hallin M, Lefevre C, Puri M L: On time-reversibility and the uniqueness of moving average representations for non-Gaussian stationary time series, Biometrika, 75(1), 170-171, (1988).

[11] Cheng Q-S: On time-reversibility of linear processes, Biometrika Vol. 86, No.2, 483-486, (1999).

[12] Tong H, Zhang Z-Q: On time-reversibility multivariate linear processes, Statistica Sinica, Vol.15, No.2, 495-504, (2005).

[13] Chan K-S, Ho L-H, Tong H: A note on time-reversibility of multivariate linear processes, Biometrika, 93(1). 221-227 (2006)

[14] MacDonald I L, Zucchini W: Hidden Markov and Other Models for Discrete-valued Time Series, London: Chapman Hall, 1997.

[15] Ball F, Milne R K, Yeo G F: Continuous-time Markov chains in a random environment, with applications to ion channel modelling. Adv. in Appl. Probab. 26 , no.4, 919–946,(1994).

[16] Churchill, G. Stochastic models for heterogeneous DNA sequences. Bull of Math Biol, 51, 79-94, (1989).

[17] Durbin R, Eddy S R , Krogh A, Mitchison G: Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids, Cambridge: Cambridge University Press, 1998.

[18] Meier A, Richardson B A, Hughes J P: Applications of continuous time hidden Markov models to the study of misclassified disease outcomes, Statistics in Medicine 22(3), 441-462, (2003).

[19] Chen Y: On the monotonicity of fluctuation spectra for three-state Markov processes, Fluctuation and Noise Letters, 7 (3) L181-192, (2007).

[20] Chiang C-L: An Introduction to Stochastic Processes and their Applications, Huntington, N.Y.: Robert E. Krieger Publishing Co., 1980.

[21] Kedem B: Binary time series, New York: Marcel Dekker, 1980.

[22] Baum L E, Petrie T: Statistical inference for probabilistic functions of finite state Markov chains, Ann. Math. Stat. 37, 1554-1563, (1966).

[23] Qian H, Elson E L, Single-molecule enzymology: stochastic Michaelis-Menten kinetics, Biophysical Chemistry, Vol. 101-102, 565-576, (2002).